SOVIET PHYSICS JETP VOLUME 32, NUMBER 4 APRIL, 1971

# INTERBAND ABSORPTION OF LIGHT IN STRONGLY DOPED SEMICONDUCTORS

B. I. SHKLOVSKIĬ and A.L. ÉFROS

I. F. Ioffe Physico-technical Institute, USSR Academy of Sciences

Submitted April 23, 1970

Zh. Eksp. Teor. Fiz. 59, 1343–1352 (October, 1970)

A theory is constructed for the tails of interband light absorption in the frequency region where the quantum energy ħω is smaller than the width of the forbidden band $E_g$. The case when there is no Fermi filling of the upper band is considered. The method employed is applicable in the region of quantum-energy deficits $\Delta = E_g - \hbar\omega$ such that the absorption coefficient $\alpha(\Delta)$ is exponentially small. It is shown that $\alpha(\Delta)$ at not too large values of Δ does not duplicate the tail of the density of states and decreases like $\exp\{-\frac{2}{5}\pi^{-1/2}(\Delta/E_B)^{5/4}(Na^3)^{-1/2}\}$. At larger Δ, the absorption duplicates the density of states calculated in[1,2].

1. The present paper is devoted to the theory of interband absorption of light with quantum energy ħω smaller than the width of the forbidden band $E_g$. In a strongly doped semiconductor, this absorption is connected with the electron levels produced in the forbidden band as a result of fluctuations of the shallow-impurity concentration. We consider the case when the extrema of the valence band and of the conduction bands are at the same point in k-space. The experimentally-observed coefficient of interband absorption $\alpha(\Delta)$ decreases exponentially with increasing quantum-energy deficit $\Delta = E_g - \hbar\omega$. The characteristic energy of this decrease increases with the impurity concentration, and its magnitude is usually smaller by two orders of magnitude than the width of the forbidden band $E_g$. We therefore always have $\Delta \ll E_g$ in the experimentally-observed region. This makes it possible to construct a theory of electronic states responsible for the absorption, within the framework of the effective-mass method.

For concreteness, we consider a semiconductor of n-type, in which, however, there are both donors and acceptors. The density of states in the forbidden band was investigated in[1,2]. We shall assume that the hole mass is larger than the electron mass. Therefore, in accordance with the terminology of[1], the case considered by us is "classical" for both hands. The density of the fluctuation levels in the forbidden band is shown dashed in Fig. 1. At not very large ϵ, the density of states decreases like

$$\ln\frac{\rho(\varepsilon)}{\rho(0)} = -\frac{\varepsilon^2}{\gamma^2}, \quad \gamma = 2\sqrt{\pi}\frac{e^2}{\varkappa r_0}(Nr_0{}^3)^{1/2}, \tag{1}$$

where ϵ is the energy reckoned towards the interior of the forbidden band, e is the electron charge, $r_0$ is the screening radius, κ is the dielectric constant, and N is the impurity concentration. In this energy region, the fundamental role is played by fluctuations of the impurity concentration, with a dimension on the order of the screening radius $r_0$; these fluctuations are described by Gaussian statistics. When $\epsilon > E_1 = \gamma^{4/3}/E_B^{1/3}$ [1], more compact clusters of the attracting impurities become significant, and form nuclei of multiply-charged atoms. The law governing the decrease of the density of states is

$$\ln\frac{\rho(\varepsilon)}{\rho(0)} = -\sqrt{\frac{\varepsilon}{E_B}}\ln\left[\left(\frac{\varepsilon}{E_B}\right)^2\frac{1}{Na^3}\right], \tag{2}$$

where $E_B$ is the ionization energy of the isolated impurity in the semi-conductor, and a is the effective Bohr radius. The density of states of the electrons in this region is determined only by the donors, and the density of the hole states only by the acceptors. We note that the mean-squared impurity potential γ in (1) includes the total concentration of the donors and acceptors.

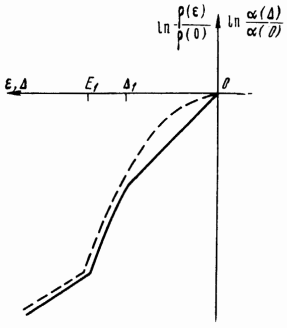

FIG. 1

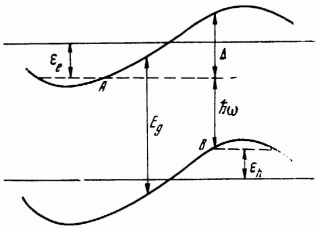

FIG. 2

FIG. 1. Logarithm of the density of states $\rho(\epsilon)$ as a function of the energy ϵ reckoned towards the interior of the forbidden band (dashed curve), and the logarithm of the coefficient of interband absorption of light, as functions of the deficit Δ (solid curve). The curves are shown for the case $r_0 < a$.

FIG. 2. Distortion of the bands by the impurity potential. The straight lines represent the bottom of the conduction band and the top of the valence band in the absence of impurities. Upper curve—potential energy of the electron in the field of the impurity-concentration fluctuation as a function of the coordinate. Lower curve—potential energy of the hole, taken with the minus sign. It is obtained from the upper curve by a downward parallel shift equal to the energy $E_g$.

The theory developed in[1-3] for the absorption of light pertains to the case when the electrons are degenerate and the compensation is small, so that the Fermi level is high in the conduction band. In this case the frequency dependence of the light-absorption coefficient has the same form as the density of states of the valence band. We consider in this paper the case when the Fermi level lies deep in the forbidden band. This may be the consequence of high temperature or of exact compensation. This pertains also to the situation (existing, for example, in germanium[4]) when a direct transition to an unfilled band takes place, and the electrons are in a lateral extremum lying lower than this band.

As noted earlier [3,4], in this case the absorption coefficient, generally speaking, does not have the same behavior as the density of states. This is clearly seen from Fig. 2. Let us assume that we investigate a transition resulting in the formation of an electron and a hole with energies $\epsilon_e$ and $\epsilon_h$ (we reckon the energies upward from the bottom of conduction band and from the top of the valence band). This is accompanied by absorption of a quantum with a deficit $\Delta = \epsilon_h - \epsilon_e$. As seen from Fig. 2, if $\Delta > 0$, then the regions in which the electron and hole are located are separated in space. Indeed, the region to the right of the point A is classically forbidden to the electron, and the region to the left of the point B is classically forbidden to the hole. The absorption coefficient contains the overlap integral of the wave functions of the electron and the hole. If the electron and hole were to be classical particles, then the absorption of the light quanta with energy smaller than $E_g$ would be impossible. At the same time, the density of states in the forbidden band would be different from zero and would satisfy formula (1), which is purely classical.

Thus, the absorption of the light is limited not only by the probability of occurrence of fluctuation levels in the forbidden band, but also by the tunneling probability, which is likewise exponentially small. Therefore the fluctuations that determine the density of states cannot be optimal for the absorption of light. Nonetheless, as shown in[2], in the case of large deficits, when formula (2) is valid, the absorption coefficient behaves like the density of states. (For the derivation given in[3] it is immaterial whether the electrons are degenerate or not.) This can be readily understood by recognizing that in this case the principal role is played by impurity clusters forming the nuclei of the atoms. To absorb a light quantum with a deficit on the order of the Bohr energy of this atom, the electron must tunnel through a distance on the order of the radius of its state. The probability of such a tunneling is not small, and therefore the absorption of the light is limited only by the number of states with energy on the order of $\Delta$. Thus, when $\Delta > E_1$ the absorption coefficient is described by the formula

$$\ln\frac{\alpha(\Delta)}{\alpha(0)} = -\sqrt{\frac{\Delta}{E_B}}\ln\left[\left(\frac{\Delta}{E_B}\right)^2\frac{1}{Na^3}\right]. \tag{3}$$

The situation is entirely different in the region of small deficits. In the region of applicability of (1), for the density of states an important role is played by large-scale fluctuations producing a smoothly varying potential. The probability of tunneling in this case is very small and such fluctuations, as shown below, are not optimal for the absorption of light. Thus, the absorption of light in this case calls for a special analysis. Essentially this is the Franz-Keldysh law, but not in an external electric field but in the random field of charged impurities. This problem was considered by Redfield and Afromovich[5]. They averaged the coefficient of light absorption in a homogeneous electric field. In the averaging they used the probability of randomly placed impurities producing a given field at a certain point. (It was assumed in fact that the entire field is produced by the nearest impurity.) It turned out that an important role in the integration is played by the field $\mathscr{E}_c = \Delta\sqrt{m\Delta}/\hbar e$, which is produced by an impurity located at a distance $r_c = a(E_B/\Delta)^{3/4}$ from the electron. It can be shown, however, that this method is applicable if the inequalities $r_c > a$, $r_c > \Delta/e\mathscr{E}_c$, and $r_c < N^{-1/3}$ are simultaneously satisfied. This is possible only in the case of weak doping ($Na^3 < 1$) and only when $\Delta < E_B$. We are interested, on the other hand, in the case $Na^3 > 1$ and $\Delta > E_B$.

2. Our main idea is to find an impurity configuration that makes the maximum contribution to the absorption of light with a given deficit. In the region of large deficits, where the absorption coefficient is exponentially small, the number of impurities participating in the fluctuation is large. Then the maximum corresponding to the sought fluctuation turns out to be very steep. Therefore the principal term $\ln\alpha$ is determined by the contribution made from the optimal fluctuation. Of course, such a method does not make it possible to determine the pre-exponential term in $\alpha(\Delta)$.

In this section we present a simplified derivation, which makes it possible to determine the argument of the exponential in $\alpha(\Delta)$ accurate to a numerical factor. This factor is determined in the rigorous theory, which will be presented in the succeeding sections.

Let us assume that fluctuation of the impurity concentration has produced a homogeneous electric field E in a volume with linear dimension R. We shall assume that R and $\mathscr{E}$ are connected by the relation $e\mathscr{E}R = \Delta$. This is necessary in order that a field of such a fluctuation be capable of absorbing a quantum with deficit $\Delta$. (The absorption of the quantum will take place also when $R > \Delta/e\mathscr{E}$, but of course such a fluctuation is less probable.) The excess number of impurities Z needed to produce the required fluctuation is determined from the condition $\mathscr{E} = Ze/\kappa R^2$, i.e.,

$$Z = \mathscr{E}\varkappa R^2/e = \varkappa R\Delta/e^2.$$

(We do not write out here the numerical factors which depend on the form of the fluctuation.) The contribution made to the absorption coefficient by such a fluctuation is proportional to

$$\exp\left(-\frac{Z^2}{NR^3}\right)\exp\left(-\frac{R\sqrt{m\Delta}}{\hbar}\right). \tag{4}$$

The first factor is the probability of the appearance of Z excess impurities in the volume $R^3$. The second factor is the probability of tunneling through a depth R. Here m is the effective mass of the lightest carrier. Expressing Z in terms of R and $\Delta$, we rewrite (4) in the form

$$\exp\left(-\frac{\varkappa^2\Delta^2}{e^4NR} - \frac{R\sqrt{m\Delta}}{\hbar}\right), \tag{5}$$

from which it is clear that at a given $\Delta$ the tunneling probability increases with decreasing R, whereas the probability of fluctuation production decreases.

We determine the dimension $R_c$ of the optimal fluctuation from the condition that the argument of the exponential in (5) be a maximum

$$R_c = a\left(\frac{\Delta}{E_B}\right)^{3/4}\frac{1}{(Na^3)^{1/2}}. \tag{6}$$

Substituting (6) in (5), we obtain the contribution made to $\alpha(\Delta)$ by the optimal fluctuation. In accordance with the foregoing, we obtain ultimately

$$\ln\frac{\alpha(\Delta)}{\alpha(0)} = -\beta\left(\frac{\Delta}{E_B}\right)^{5/4}\frac{1}{(Na^3)^{1/2}}. \tag{7}$$

Here $\beta$ is the numerical factor which cannot be obtained from such a simplified derivation. As will be shown below, $\beta = 2/5\sqrt{\pi}$. It is seen from (7) that the law governing the decrease of the absorption coefficient differs significantly from the density of states (1). We note that (7), like (1), contains the total donor and acceptor concentration.

Let us now investigate the region of applicability of (7). We did not take into account in the foregoing derivation the fact that the Coulomb potential of the impurities is screened by the electrons. Therefore the result is valid if $R_C < r_0$, i.e., if $\Delta < \Delta_1$, where $\Delta_1 = E_B(r_0/a)^{4/3}(Na^3)^{2/3}$. If $R_C > r_0$, then when $R = r_0$ the second term in (5) is small compared with the first, i.e., the tunneling is insignificant. The probability of producing a potential well with depth $\Delta$ and dimension R with $R > r_0$ decreases with increasing R[1]. (The potential of the remote impurities does not reach the center.) Therefore, just as in the theory of the density of states, the dimension of the optimal cluster is in this case $r_0$. Substituting $R = r_0$ in (5), we obtain for $\Delta > \Delta_1$

$$\ln\frac{\alpha(\Delta)}{\alpha(0)} = -\frac{\Delta^2}{\gamma^2}, \tag{8}$$

which varies just like the density of states.

If $\Delta_1 < E_1$, i.e., $r_0 < a$, then with increasing $\Delta$ the absorption coefficient passes through three regions. Formula (7) holds when $\Delta < \Delta_1$, formula (8) when $E_1 > \Delta > \Delta_1$, and formula (3) when $\Delta > E_1$. This situation is illustrated by the solid curve of Fig. 1.

If $\Delta_1 > E_1$ ($r_0 > a$), then fluctuations of the type of the atom become significant before $R_C$ becomes equal to $r_0$. Indeed, comparing (3) and (7), we see that this occurs at $\Delta = \Delta_2 = E_B(Na^3)^{2/3} < E_1, \Delta_1$. Thus, in the case when $r_0 > a$, Eq. (7) is valid when $\Delta < \Delta_2$, Eq. (3) when $\Delta > \Delta_2$, and there is no region where (a) is valid. We recall that in the spirit of our derivation, formulas (3), (7), and (8) are valid only if

$$\left|\ln\frac{\alpha(\Delta)}{\alpha(0)}\right| \gg 1.$$

Taking this circumstance into account and bearing in mind $\Delta \gg E_B$, we can easily show that a region of applicability of (7) and (8) exists only when $Na^3 \gg 1$.

We note further that the fluctuation in question, as seen from Fig. 2, represents a potential well for the electron and an adjacent potential well for the hole. In our derivation it is assumed, in essence, that states with binding energies on the order of the well depths $\Delta$ exist in these wells. The validity of this assumption can be verified by recognizing that the well dimension is of the order of $R_C$, and consequently the energy of the ground state is located at a distance $\hbar^2/mR_C^2 \ll \Delta$ away from the bottom of the well. The latter inequality follows from the fact that the logarithm in (7) is always much larger than unity. It can thus be assumed that in the wells considered here the levels are very close. This circumstance will play an important role in the derivation of the rigorous theory.

Thus, the main feature of the interband absorption in the absence of filling of the upper band is that there exists a region in which the absorption coefficient does not behave like the density of states, and $\ln\alpha(\Delta)$ decreases in accordance with (7). In the experimentally observed region, it is apparently difficult to distinguish between this decrease and a linear one.

Let us now discuss the possibility of experimentally observing (7). The region of existence of (7) is large if the screening radius $r_0$ is large. This would be aided by a high temperature and strong compensation[1)]. When the concentration changes from $Na^3 \ll 1$ to $Na^3 \gg 1$, a transition from (3) to (7) takes place. This can be revealed by the dependence of the characteristic energy of the decrease (slope) of the function $\ln\alpha(\Delta)$ on N. This dependence is logarithmic at small N and of the square-root type at large N. The experimental data[4] pertaining to direct transitions in germanium do not contradict our conclusions. The order of magnitude of the slopes and their dependence on N are similar to those expected. It was difficult to draw more definite conclusions in view of the large errors in the determination of the slopes and of the impurity concentration.

3. The next two sections are devoted to a rigorous derivation of (7) and to a determination of the coefficient $\beta$. The wave function of the electron in the classically inadmissible region changes significantly over a distance $\hbar/\sqrt{m\Delta}$. This distance is large compared with the average distance between the impurities $N^{-1/3}$ in the entire region of applicability of (7). Therefore the fluctuations in question can be characterized by a macroscopic concentration of the excess impurities $\xi(\mathbf{r})$, which is obtained from the excess microscopic impurity density $N(\mathbf{r}) - N$ by averaging over distances that are large compared with the average distance between the impurities, but are small compared with the characteristic dimension of the variation of the wave function. The potential energy of the electron $U_\mathbf{r}\{\xi\}$ in the field of the fluctuation $\xi(\mathbf{r})$, at distances small compared with the screening radius $r_0$, is given by

$$U_\mathbf{r}\{\xi\} = -\frac{e^2}{\varkappa}\int\frac{\xi(\mathbf{r}')\,d^3r'}{|\mathbf{r}-\mathbf{r}'|}. \tag{9}$$

The absorption coefficient $\alpha(\Delta)$ can be presented in the form of a functional integral over all the functions $\xi(\mathbf{r})$:

$$\alpha(\Delta) = \int D\xi\,\alpha_\Delta\{\xi\}\,e^{-\Omega\{\xi\}}. \tag{10}$$

Here $\exp(-\Omega)$ is the probability of producing the fluctuation $\xi(\mathbf{r})$. If the impurities have a random distribution[6], then

$$\Omega = \int\left[(N+\xi)\ln\frac{(N+\xi)}{N} - \xi\right]d^3r. \tag{11}$$

As already mentioned, we are interested in the

---

[1)] We note, however, that in the case of appreciable compensation, the role of the screening radius may be assumed by the characteristic length of the correlation in the impurity location.

region of Gaussian statistics, where $\xi \ll N$. In this case

$$\Omega = \int \frac{\xi^2}{2N} d^3r. \tag{12}$$

The functional $\alpha_\Delta\{\xi\}$ is the coefficient of absorption with an energy deficit $\Delta$ in the field of the fluctuation $\xi(\mathbf{r})$

$$\alpha_\Delta\{\xi\} = C \sum_{mn} P_{mn}\{\xi\} \delta(\varepsilon_h{}^m - \varepsilon_e{}^n - \Delta), \tag{13}$$

where C is a quantity independent of $\Delta$, $\epsilon_h^m$ and $\epsilon_e^n$ are the energies of the hole and electron states (see Fig. 2), $P_{mn}$ is the square of the modulus of the overlap integral

$$P_{mn}\{\xi\} = \left| \int \Psi_h{}^m(\mathbf{r}) \Psi_e{}^n(\mathbf{r}) d^3r \right|^2, \tag{14}$$

and $\Psi_h^m$ and $\Psi_e^n$ are the wave functions of the hole and of the electron. As seen from Fig. 2, at large $\Delta$ the regions that are classically admissible for the electron and the hole are separated in space and therefore $P_{mn}$ is small.

We shall assume henceforth that the mass of the hole is much larger than the mass of the electron. This allows us to assume that the hole does not tunnel beyond the surface $U_\mathbf{r}\{\xi\} = \epsilon_h^m$ (we call this the h surface), and consequently the integral in (14) is determined by the value of the wave function of the electron on this surface. Actually, we need only the argument of the exponential of the wave function (the imaginary part of the phase). To find it, we use a quasi-classical method, the applicability condition of which is

$$\frac{d}{dr} \left| \frac{\hbar}{\sqrt{2m(\varepsilon_e{}^n - U(\mathbf{r}))}} \right| \ll 1. \tag{15}$$

As is clear from the statements made in Sec. 2, this condition reduces to the inequality $\hbar/\sqrt{m\Delta}R_c \ll 1$, which is always satisfied in the region where (7) is valid.

We shall seek the wave function of the electron in the classically forbidden region in the form

$$\Psi = A(\mathbf{r}) e^{iS(\mathbf{r})/\hbar}, \tag{16}$$

where the truncated action S satisfies the Hamilton-Jacobi equation. Assume that we know the wave function on the surface $\epsilon_e^n = U_\mathbf{r}\{\xi\}$ (the e surface). By the same token, the action $S_0(\mathbf{r})$ and its gradient $\nabla S$ are specified at each point of the surface. The normal component of the gradient is determined from the Hamilton-Jacobi equation, which yields on the e surface $(\nabla S)^2 = 0$. We draw for each point of the e surface trajectories (rays) satisfying the classical equations of motion and havong on the e surface a momentum $\mathbf{p} = \nabla S$. The action at an arbitrary point $\mathbf{r}$ can be represented in the form of an integral over the particular trajectory that passes through the point $\mathbf{r}$:

$$S(\mathbf{r}) = S_0(\mathbf{r}_e) + \int_{\mathbf{r}_e}^{\mathbf{r}} \sqrt{2m(\varepsilon_e{}^n - U(\mathbf{r}))}\, dl, \quad U(\mathbf{r}) \equiv U_\mathbf{r}\{\xi\}, \tag{17}$$

where $\mathbf{r}_e$ is the point at which this trajectory intersects the e surface and $dl$ is an element of the trajectory length[2]. If $\mathbf{r}$ is in a classically forbidden region, then the trajectory and the point $\mathbf{r}_e$ may turn out to be complex[7]:

We shall show that the states of importance are those close to the bottom of the well, i.e., with kinetic energy $K \ll \Delta$ (of course, $K \gtrsim \hbar^2/mR_c^2$). On the e surface, the components of $\nabla S$ are of the order of $\sqrt{mK}$. On the h surface we have $|\nabla S| \sim \sqrt{m\Delta} \gg \sqrt{mK}$, and therefore when solving the mechanical problem it can be assumed that the momentum on the e surface is equal to zero. It can easily be shown that in this case the trajectories in the classically forbidden region are real. (The time characterizing motion along these trajectories is imaginary.) These trajectories can be obtained by solving the mechanical problem of motion with potential energy $-U(\mathbf{r})$, total energy $-\epsilon$, and zero initial velocity.

In the three-dimensional case, the wave function on the e surface may turn out to be exponentially small, i.e., $\hbar^{-1} \operatorname{Im} S_0(\mathbf{r}) > 1$. (Let us imagine, for example, that the potential energy depends on one coordinate, and the particle has a momentum along another coordinate.) However, the maximum imaginary momentum on the e surface is $\sim \sqrt{mK}$, and on the h surface of the order of $\sqrt{m\Delta}$. We can therefore neglect the first term of (17). Let us determine the argument of the exponential in the integral (14). To this end, we draw from all the points of the h surface trajectories satisfying the equations of motion and intersecting the e surface with zero velocity. We furthermore find an h-surface point $\mathbf{r}_h$ for which the pure imaginary integral in (17) is minimal. The integrand in (14) has a sharp maximum at this point, and therefore

$$P_{nm} \sim \exp\left\{ -\frac{2}{\hbar} \operatorname{Im} S(\mathbf{r}_h) \right\}. \tag{18}$$

In similar fashion we obtain the principal term of the sum (13). To this end we examine all the h and e surfaces (with potential energies differing by $\Delta$) for a given potential and find points $\tilde{\mathbf{r}}_e$ and $\tilde{\mathbf{r}}_h$ $(U(\tilde{\mathbf{r}}_e) - U(\tilde{\mathbf{r}}_h) = -\Delta)$ such that the action calculated along the trajectory emerging from $\tilde{\mathbf{r}}_h$ and arriving from $\tilde{\mathbf{r}}_e$ with zero momentum is minimal. Then, retaining in (13) only the principal term of the sum, we obtain

$$\alpha_\Delta\{\xi\} = \exp(-\Phi\{\xi\}), \tag{19}$$

where

$$\Phi\{\xi\} = \frac{2}{\hbar} \int_{\tilde{\mathbf{r}}_e}^{\tilde{\mathbf{r}}_h} \sqrt{2m(U_\mathbf{r}\{\xi\} + \Delta - U_{\tilde{\mathbf{r}}_h}\{\xi\})}\, dl, \tag{20}$$

with

$$U_{\tilde{\mathbf{r}}_h}\{\xi\} - U_{\tilde{\mathbf{r}}_e}\{\xi\} = \Delta. \tag{20a}$$

Since the fluctuation in question is macroscopic, the integrand in (10) has a sharp maximum at a certain function $\bar{\xi}(\mathbf{r})$. Therefore, in accordance with the statements made in Sec. 2, we have

$$\ln \frac{\alpha(\Delta)}{\alpha(0)} = -\Omega\{\bar{\xi}\} - \Phi\{\bar{\xi}\}, \tag{21}$$

---

[2] Strictly speaking, the quasiclassical approximation is not valid on the surface $\epsilon_e^n + U(\mathbf{r})$. Our arguments, however, remain unchanged if we take the e surface to mean the surface $\epsilon_e^n + T = U(\mathbf{r})$, where the energy T is chosen such as to satisfy the condition (15) on the surface, but $T \ll \Delta$.

where $\tilde{\xi}$ is a function that minimizes the functional $\Omega + \Phi$. To find this function we calculate first the variation of $\Phi\{\xi\}$. In (20), $\tilde{\mathbf{r}}_h$ is a functional of $\xi$, and the prescription for constructing this functional was given above. The vector $\tilde{\mathbf{r}}_e$ is a functional of $\xi$ and a function of $\tilde{\mathbf{r}}_h$. Finally, the potential energy and the form of the trajectory are functionals of $\xi$.

We note first that when (20) is varied $\tilde{\mathbf{r}}_h$ must be regarded as constant, since, in accordance with the method used for its construction, the functional $\Phi$ increases for an arbitrarily small change of $\tilde{\mathbf{r}}_h$, i.e., $(\partial\Phi/\partial\tilde{\mathbf{r}}_h)\delta\tilde{\mathbf{r}}_h = 0$. The variation of (20) with respect to the lower limit is equal to zero, since the corresponding derivative is equal to the momentum at the point $\tilde{\mathbf{r}}_e$. Finally, the change of $\Phi$ due to the variation of the form of the trajectory is equal to zero, as a result of the Maupertuis principle[8]. Thus, in the variation of (20) it is necessary to take into account only the explicit functional dependence of $U_{\mathbf{r}}\{\xi\}$. This yields

$$\delta\Phi\{\xi\} = \frac{\sqrt{2m}}{\hbar}\int_{\tilde{\mathbf{r}}_e}^{\tilde{\mathbf{r}}_h} \frac{(\delta U_{\mathbf{r}}\{\xi\} - \delta U_{\tilde{\mathbf{r}}_h}\{\xi\})\,dl}{\sqrt{U_{\mathbf{r}}\{\xi\} - U_{\tilde{\mathbf{r}}_h}\{\xi\} + \Delta}}, \tag{22}$$

where, according to (9),

$$\delta U_{\mathbf{r}}\{\xi\} = -\frac{e^2}{\varkappa}\int \frac{\delta\xi(\mathbf{r}')\,d^3r'}{|\mathbf{r} - \mathbf{r}'|}. \tag{23}$$

Substituting (23) in (22), we obtain

$$\delta\Phi = \frac{e^2\sqrt{2m}}{\varkappa\hbar}\int d^3r\,\delta\xi(\mathbf{r}) \int_{\tilde{\mathbf{r}}_e}^{\tilde{\mathbf{r}}_h} \left(\frac{1}{|\tilde{\mathbf{r}}_h - \mathbf{r}|} - \frac{1}{|\mathbf{r}' - \mathbf{r}|}\right) \cdot \frac{dl'}{\sqrt{U(\mathbf{r}') + \Delta - U(\tilde{\mathbf{r}}_h)}}. \tag{24}$$

From (12) we have

$$\delta\Omega = \int \frac{\xi\delta\xi}{N}\,d^3r. \tag{25}$$

From the condition $\delta\Omega + \delta\Phi = 0$ we obtain equations for the determination of the extremal function $\tilde{\xi}(\mathbf{r})$:

$$\tilde{\xi}(\mathbf{r}) = \frac{Ne^2\sqrt{2m}}{\varkappa\hbar}\int_{\tilde{\mathbf{r}}_e}^{\tilde{\mathbf{r}}_h} \left(\frac{1}{|\mathbf{r}' - \mathbf{r}|} - \frac{1}{|\tilde{\mathbf{r}}_h - \mathbf{r}|}\right) \frac{dl'}{\sqrt{U(\mathbf{r}') + \Delta - U(\tilde{\mathbf{r}}_h)}}, \tag{26}$$

where $U(\mathbf{r})$ is expressed in terms of $\tilde{\xi}$ by means of formula (9), and the points $\tilde{\mathbf{r}}_e$ and $\tilde{\mathbf{r}}_h$ are connected by the condition (28) and by the classical equation for the trajectory[8]

$$\frac{d^2\mathbf{r}}{dl^2} = \left[\frac{d\mathbf{r}}{dl}\left(\nabla U \frac{d\mathbf{r}}{dl}\right) - \nabla U\right] \Big/ 2(U(\tilde{\mathbf{r}}_e) - U(\mathbf{r})), \tag{27}$$

the velocity at the point $\tilde{\mathbf{r}}_e$ being equal to zero.

We introduce the dimensionless quantities

$$\mathbf{x} = \frac{\mathbf{r}}{R_c}, \quad R_c = a\left(\frac{\Delta}{E_B}\right)^{3/4} \frac{1}{(Na^3)^{1/2}},$$

$$\mathrm{F}(\mathbf{x}) = \frac{\tilde{\xi}}{N}\left(\frac{\Delta}{E_B}\right)^{1/2}, \quad W(\mathbf{x}) = \frac{U}{\Delta}. \tag{28}$$

In terms of this notation, Eqs. (26), (9), and (20a) take the form

$$F(\mathbf{x}) = 2\int_{\mathbf{x}_2}^{\mathbf{x}_1} \left(\frac{1}{|\mathbf{x}' - \mathbf{x}|} - \frac{1}{|\mathbf{x}_1 - \mathbf{x}|}\right) \frac{ds'}{\sqrt{1 + W(\mathbf{x}') - W(\mathbf{x}_1)}} \tag{29}$$

$$W(\mathbf{x}) = -2\int \frac{F(\mathbf{x}')\,d^3x'}{|\mathbf{x} - \mathbf{x}'|}, \tag{30}$$

$$W(\mathbf{x}_1) - W(\mathbf{x}_2) = 1. \tag{31}$$

Here $\mathbf{x}_1 = \tilde{\mathbf{r}}_h/R_c$, $\mathbf{x}_2 = \tilde{\mathbf{r}}_e/R_c$, and the trajectory $s(\mathbf{x})$ is determined by Eq. (27), which in terms of the new variables likewise does not contain any parameters. Substituting (28), (12), and (20) in (21), we obtain the result (7), where

$$\beta = 2\int_{\mathbf{x}_2}^{\mathbf{x}_1} [W(\mathbf{x}) + 1 - W(\mathbf{x}_1)]^{1/2}\,ds + \frac{1}{2}\int F^2(\mathbf{x})\,d^3x, \tag{32}$$

and $F(\mathbf{x})$ and $W(\mathbf{x})$ constitute the solution of the dimensionless equations (29)–(31).

4. We shall show that the solution of Eqs. (27), (29)–(31) is a straight-line trajectory joining the points $\mathbf{x}_1$ and $\mathbf{x}_2$. Let us assume that this is so. Then it follows from (29) that the distribution of the concentration $F(\mathbf{x})$, and consequently also of the potential $W(\mathbf{x})$ has cylindrical symmetry with respect to this straight line. This means that at any point of the trajectory the force is directed along the trajectory. It follows therefore that such a trajectory satisfies the equations of motion, since at the point $\mathbf{x}_2$ the particle velocity is equal to zero.

We introduce cylindrical coordinates $\eta$ and $\rho$ with origin at the point $\mathbf{x}_2$ and with the axis $\eta$ directed along the trajectory. We then obtain from (29)

$$F(\rho, \eta) = -2\int_0^{\eta_1} \left\{\frac{1}{[(\eta_1 - \eta)^2 + \rho^2]^{1/2}} - \frac{1}{[(\eta' - \eta)^2 + \rho^2]^{1/2}}\right\} \frac{d\eta'}{\sqrt{W(\eta') - W(0)}}. \tag{33}$$

Substituting (33) in (30), we obtain an equation for the potential on the $\eta$ axis

$$W(\eta) = 8\pi\int_0^{\eta_1} (|\eta' - \eta| - |\eta_1 - \eta|)\frac{d\eta'}{\sqrt{W(\eta') - W(0)}}. \tag{34}$$

It follows from (34) that when $\eta < 0$

$$W(\eta) = W(0) = 8\pi\int_0^{\eta_1} (\eta' - \eta_1)\frac{d\eta'}{\sqrt{W(\eta') - W(0)}}, \tag{35}$$

and when $\eta > \eta_1$ we have $W(\eta) = -W(0)$. In the region $0 < \eta < \eta_1$, differentiating (34) twice with respect to $\eta$, we obtain the equation

$$\frac{d^2V}{d\eta^2} = \frac{16\pi}{\sqrt{V(\eta)}} \tag{36}$$

with boundary conditions $V(0) = 0$; $dV/d\eta\,|_{\eta=0} = 0$. Here, by definition, $V(\eta) = W(\eta) - W(0)$. The solution of (36) is

$$V(\eta) = (6\sqrt{\pi}\eta)^{4/3}, \quad 0 < \eta < \eta_1. \tag{37}$$

From (31) we find that $\eta_1$ is determined by the condition $V(\eta_1) = 1$, i.e., $\eta_1 = \frac{1}{6}\sqrt{\pi}$. Substituting (37) in (35), we get $W(0) = -\frac{1}{2}$.

Thus, the potential of the optimal fluctuation $U(\eta)$ varies along the $\eta$ axis in the following manner: when $\eta < 0$ it is constant and equals $-\Delta/2$, when $\eta > \eta_1$ it is equal to $\Delta/2$, and in the intermediate region its variation is given by (37). Of course, if screening is taken into account, then the potential on the axis vanishes at distances larger than $r_0$. The obtained potential satisfies the assumptions made concerning the smoothness of its variation. Therefore there exist in such a potential states with small kinetic energy. In addition, the form of the optimal potential is such that absorption of a quantum with energy Δ is possible only as a result of these states.

The charge distribution in the optimal fluctuation can be obtained by substituting (37) in (33). As $x \to \infty$, the function $F(x)$ decreases like $1/x^2$. Therefore the second interval in (32) converges. After rather laborious calculations we obtain $\beta = 2/5\sqrt{\pi}$ from (32), (33), and (37).

---

[1] B. I. Shklovskiĭ and A. L. Éfros, Fiz. Tekh. Poluprov. 4, 305 (1970) [Sov. Phys.-Semicond. 4, 249 (1970).

[2] B. I. Shklovskiĭ and A. L. Éfros, Zh. Eksp. Teor. Fiz. 58, 657 (1970) [Sov. Phys.-JETP 31, 351 (1970)].

[3] M. I. Dyakonov, A. L. Efros, and D. L. Mitchell, Phys. Rev., 180, 819 (1969).

[4] L. V. Keldysh and G. P. Proshko, Fiz. Tverd. Tela 5, 3378 (1963) [Sov. Phys.-Solid State 5, 2481 (1964)].

[5] D. Redfield, Phys. Rev. 130, 916 (1963); M. A. Afromovich and D. Redfield, Trudy IX Mezhdunarodnoĭ konferentsii po fizike poluprovodnikov (Transactions of Ninth International Conference on Semiconductor Physics), 1, Nauka, L. 1969, p. 103.

[6] I. M. Lifshits, Zh. Eksp. Teor. Fiz. 53, 743 (1967) [Sov. Phys.-JETP 26, 462 (1968)].

[7] Yu. A. Kravtsov, Izvestiya vyssh. uch. zaved., Radiofizika, 10, 1283 (1967).

[8] L. D. Landau and E. M. Lifshitz, Mekhanika (Mechanics), Fizmatgiz, 1958, p. 176 [Addison-Wesley].

Translated by J. G. Adashko
154